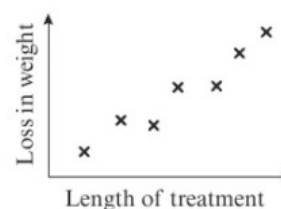


## Bivariate Data Worksheet

1.

Some research was done into the effectiveness of a weight-reducing drug. Seven people recorded their weight loss and this was compared with the length of time for which they had been treated. A scatter diagram was drawn to represent this data.

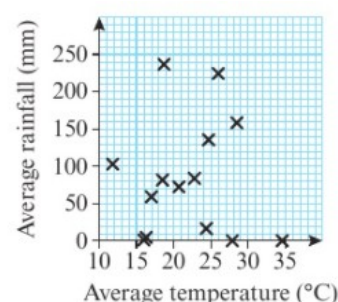


- Describe the type of correlation shown by the scatter diagram.
- Interpret the correlation in context.

2.

The average temperature and rainfall were collected for a number of cities around the world.

The scatter diagram shows this information.



- Describe the correlation between average temperature and average rainfall.
- Comment on the claim that hotter cities have less rainfall.

3.

The table below shows the packing times for a particular employee for a random sample of orders in a mail order company.

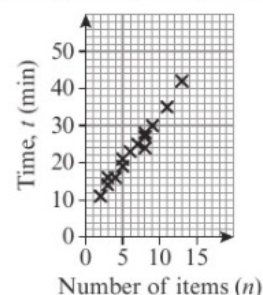
Number of items ( $n$ )	2	3	3	4	5	5	6	7	8	8	8	9	11	13
Time ( $t$ min)	11	14	16	16	19	21	23	25	24	27	28	30	35	42

A scatter diagram was drawn to represent the data.

- Describe the correlation between number of items packed and time taken. (1 mark)

The equation of the regression line of  $t$  on  $n$  is  $t = 6.3 + 2.64n$ .

- Give an interpretation of the value 2.64. (1 mark)



4.

The table shows average monthly temperature,  $t$  (°C), and the number of pairs of gloves,  $g$ , a shop sells each month.

$t$ (°C)	6	6	50	10	13	16	18	19	16	12	9	7
$g$	81	58	50	42	19	21	4	2	20	33	58	65

The following statistics were calculated for the data on temperature:

mean = 15.2, standard deviation = 11.4

An outlier is an observation which lies  $\pm 2$  standard deviations from the mean.

- Show that  $t = 50$  is an outlier. (1 mark)
- Give a reason whether or not this outlier should be omitted from the data. (1 mark)

The equation of the regression line of  $g$  on  $t$  for the remaining data is  $g = 99.6 - 5.2t$ .

- Give an interpretation of the value  $-5.2$  in this regression equation. (1 mark)

5.

James placed different masses ( $m$ ) on a spring and measured the resulting length of the spring ( $s$ ) in centimetres. The smallest mass was 20 g and the largest mass was 100 g.

He found the equation of the regression line of  $s$  on  $m$  to be  $s = 44 + 0.2m$ .

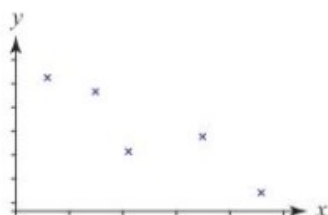
**a** Interpret the values 44 and 0.2 in this context. (2 marks)

**b** Explain why it would not be sensible to use the regression equation to work out:

- i** the value of  $s$  when  $m = 150$       **ii** the value of  $m$  when  $s = 60$ . (2 marks)

6.

This diagram relates two variables,  $x$  and  $y$



**a** Describe the correlation.

**b** What is the most likely value of  $r$ : +0.6, -0.2 or -0.7

7.

For each of these examples of bivariate data, state whether the correlation is likely to be positive, negative or zero. If non-zero, state whether you think there is a causal relationship between the variables.

- a** Daytime temperature at a seaside resort and number of deckchairs hired out.
- b** Plant growth and amount of fertilizer applied.
- c** A person's annual income and weight.
- d** Unemployment rates and measures of the standard of living.

8.

According to Department for Transport statistics, in the period 2006 - 2016 the numbers of licensed cars and vans in Britain have steadily increased whilst the number of licensed buses has steadily decreased.



- a** State the sign of the correlation between the number of cars and
  - i** vans,      **ii** buses,      **iii** the year.
- b** Explain what is meant by the statement 'the correlation between the numbers of cars and vans is an example of a spurious correlation.'